# Tutoriat5: Continuous Random Variables

Mihai Duzi    Lucan Cristian    Luminaru Ionuț

## Contents

# 1 Continuous Random Variables

## 1.1 Example

For a uniformly distributed interval $(0, 1)$, what is the probability of picking $\frac{1}{2}$? But of picking something larger than $\frac{1}{2}$?

$$P(X = \tfrac{1}{2}) = \frac{1}{\infty} = 0$$

$$P(X > \tfrac{1}{2}) = P\left(X \in \left(\tfrac{1}{2}, 1\right)\right) = \frac{1}{2}$$

## 1.2 Probability Density Function

This is how the distribution of a discrete random variable usually looks (segmented/sampled):

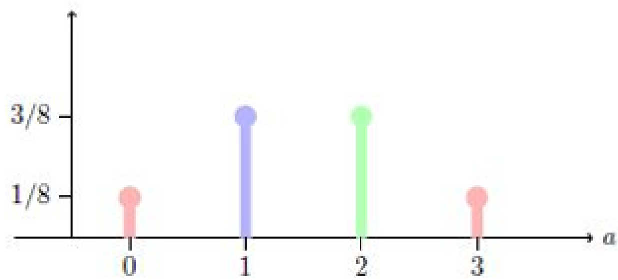


Figure 1: Discrete distribution pmf

Here, $x$ is the value and $y$ is the probability of $x$. We can directly compute $P(X = k)$.
This is how a continuous distribution usually looks (all points connected):

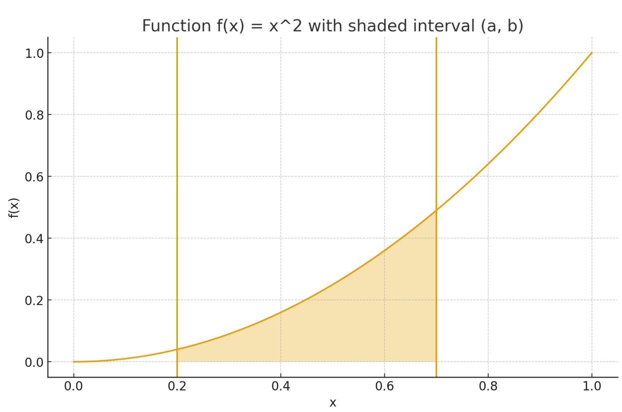


Figure 2: Continuous distribution pdf

In this scenario, we cannot pick a point and assign probability, but we can pick intervals. The analogue of the pmf is the **probability density function** $f(x)$.

To calculate the probability for the interval $(a, b)$:

$$\int_a^b f(x)\,dx$$

### 1.2.1 Properties

For a pdf $f(x)$:

1. $f(x) \geq 0$
2. $$P(\Omega) = \int_{-\infty}^{\infty} f(x)\,dx = 1$$

**Note**: If $f(x)$ is defined on $(a, b)$, then $f(x) = 0$ for $x \notin (a, b)$.

The main differences from discrete distributions: we use integrals instead of sums, and $f(x)$ does not need to be $\leq 1$.

## 1.3 Cumulative Distribution Function

This is how the cdf looks for the discrete variable:

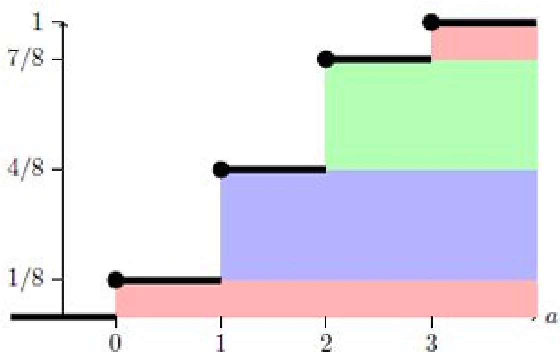


Figure 3: Discrete distribution cdf

And this is the cdf of the continuous distribution:

$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

The cdf is also called the repartition function.

We also have the tail probability which is the complementary of the cdf

$P(X \geq x) = \int_x^{\infty} f(t)dt = 1 - F(x)$

### 1.3.1 Properties

Same as discrete, with additions:

1. $P(a \leq X \leq b) = F(b) - F(a)$
2. $F'(x) = f(x)$

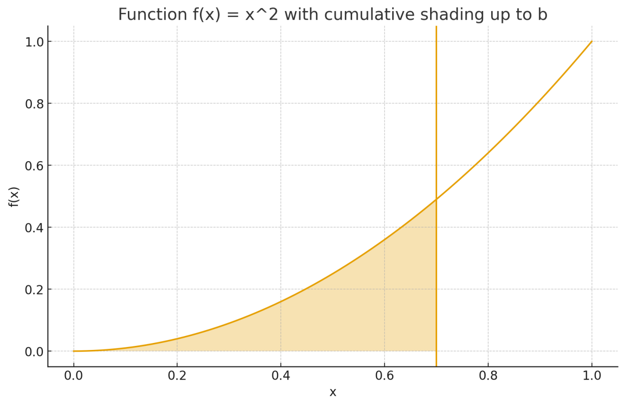


Figure 4: Continuous distribution cdf

## 1.4 Exercises

### Exercise 1

For a continuous random variable $X$ taking values in $(0, 1)$, with pdf $f(x) = Cx^2$, calculate:

a) $C = ?$

b) $F(x) = ?$

c) $P(0.3 \leq x \leq 0.7) = ?$

### Solution

a) Using the pdf property:

$$\int_0^1 Cx^2 \, dx = 1$$

$$\frac{C}{3} x^3 \Big|_0^1 = \frac{C}{3} = 1 \quad \Rightarrow \quad C = 3$$

Thus $f(x) = 3x^2$.

b) The cdf:

$$F(a) = \int_0^a 3x^2 \, dx = a^3$$

Cases:

1. $a < 0 \Rightarrow F(a) = 0$
2. $a \in (0, 1) \Rightarrow F(a) = a^3$
3. $a > 1 \Rightarrow F(a) = 1$

c)

$$P(0.3 \leq x \leq 0.7) = F(0.7) - F(0.3)$$

$$= 0.7^3 - 0.3^3 = 0.316$$

The graphs for $f(x)$ and $F(x)$:

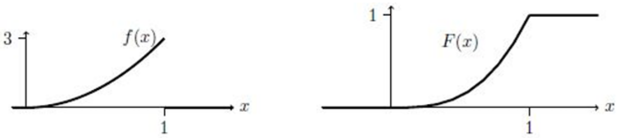


Figure 5: Exercise 1: $f(x)$ and $F(x)$

# 2 Classic distributions

## 2.1 Uniform distribution

This is the distribution you mostly used throughout your labs to generate numbers between (0, 1) to simulate experiments.

**Notation**: U(a, b), where (a, b) is the interval from which samples will be selected with the same probability.

**pdf:** $f(x) = \frac{1}{b-a}$

**cdf:** $F(x) = \frac{x-a}{b-a}$

**Visualisation:**

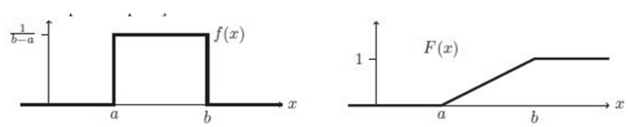


Figure 6: Example of Uniform random variable pdf and pmf

## 2.2 Exponential Distribution

This distribution models waiting times until a certain event will happen (like waiting for the ambulance to come).

**Notation**: $exp(\lambda)$, defined on $(0, \infty)$, where $\lambda$ represented the rate of occurance.

**pdf:** $f(x) = \lambda e^{-\lambda x}$

**cdf:** $F(x) = 1 - e^{-x\lambda}$

**Visualisation:**

**Example:** What is the probability of having to wait less then 7 minutes for a bus, taking into account there's two buses I can take (B1, B2, they are independent), on average having 10 buses B1 each hour and 6 buses B2 each hour? ?

First bus is modeled by $X_1 = exp(\lambda_1)$ and the second by $X_2 = exp(\lambda_2)$

The probability of no bus coming in 7 minutes is:

$P(X_1 > 7, X_2 > 7) = P(X_1 > 7)P(X_2 > 7) = e^{-\lambda_1 7}e^{-\lambda_2 7} = e^{-7(\lambda_1+\lambda_2)}$

So the probability of any coming in 7 minutes is:

$1 - e^{-7(\lambda_1+\lambda_2)}$

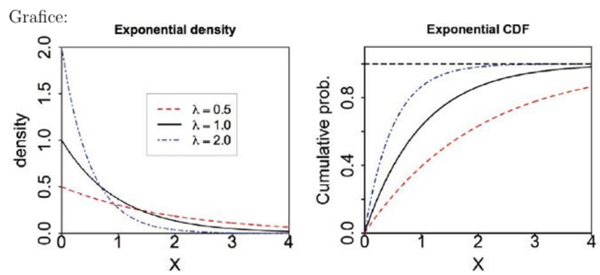


Figure 7: Example of Exponential random variable pdf and cdf

Let's calculate the $\lambda$ now
$\lambda_1 = \frac{60}{10} = 6$
$\lambda_2 = \frac{60}{6} = 10$
Pop it into the formula and you get the result

### 2.2.1 Memorylessness

Let's say I'm only waiting for the first bus (B1) now. Knowing that I waited 5 minutes already without the bus coming, what is the probability that I will have to wait at least 3 more minutes?
$P(X \geq 3 + 5 \mid X \geq 5) = \frac{P(X \geq 8 \cap X \geq 5)}{P(X \geq 5)} = \frac{P(X \geq 8)}{P(X \geq 5)} = \frac{e^{-8\lambda}}{e^{-5\lambda}} = e^{-3\lambda} = P(X \geq 3)$
**Note:** As you might have guesses by now, the exponential is the continuous representation of the geometric distribution

## 2.3 Normal distribution

The most important continuous distribution. Will appear again as we go towards statistics.
**Notation:** N$(\mu, \sigma^2)$ is a continuous random variable that is defined by it's expected value ($\mu$) and standard deviation ($\sigma$)
**pdf:** $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$
**cdf:** $e^{x^2}$ does not have a primitive so we don't know who F(x) is. For that we use tables
**Visualization:**

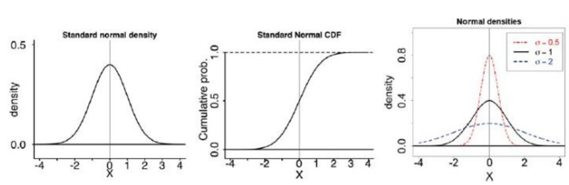


Figure 8: Different normal distributions

Z = N(0, 1) is called the standard normal distribution, having $\phi(x)$ as the pdf and $\Phi(x)$ as the cdf.
**Visualization:**
Those above probabilities also present **the big thumb rule**.

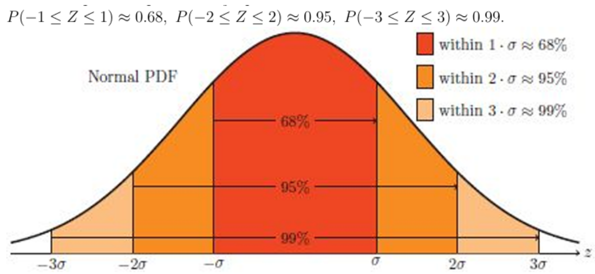


Figure 9: Normal distribution with big thumb rule

You might also notice that the probabilities are symmetric with respect to 0.

**Exercise:**

What is the value of $\Phi(1)$ ?

**Solution**

$\Phi(1) = P(Z \leq 1) = P(Z \leq -1) + P(-1 \leq Z \leq 1)$

We know the ladder from the big thumb rule, but how do we find the first?

As I said, the variable is symmetric with respect to 0, which means $P(Z \leq -1) = P(Z \geq 1)$

$P(\Omega) = P(Z \leq -1) + P(-1 \leq Z \leq 1) + P(Z \geq 1) = 2P(Z \leq -1) + P(-1 \leq Z \leq 1) = 1$

We already know the value of $P(-1 \leq Z \leq 1)$ so we can get $2P(Z \leq -1)$, and so we have everything we need to make the calculation

## 2.4 Variable substitution

### 2.4.1 Uniform distribution example

Let $X \sim U(0,2)$, so the density is $f_X(x) = 1/2$ and the cdf is $F_X(x) = x/2$ on the interval $[0,2]$.

What are the domain of values, pdf, and cdf of $Y = X^2$?

**Answer:** The domain of values of $Y$ is $[0,4]$.

To find the cdf, we use the definition:

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = F_X(\sqrt{y}) = \frac{\sqrt{y}}{2}.$$

To find the pdf, we differentiate the cdf:

$$f_Y(y) = F_Y'(y) = \left(\frac{\sqrt{y}}{2}\right)' = \frac{1}{4\sqrt{y}}.$$

### 2.4.2 Exponential distribution example

Let $X \sim \exp(\lambda)$, so $f_X(x) = \lambda e^{-\lambda x}$ on $[0, \infty)$. What is the density of $Y = X^2$?

**Answer:** We use the change of variables:

$$y = x^2 \Rightarrow dy = 2x\,dx \Rightarrow dx = \frac{dy}{2\sqrt{y}}.$$

Thus,

$$f_X(x)dx = \lambda e^{-\lambda x}dx = \lambda e^{-\lambda\sqrt{y}}\frac{dy}{2\sqrt{y}} = f_Y(y)dy.$$

Therefore,

$$f_Y(y) = \frac{\lambda}{2\sqrt{y}}e^{-\lambda\sqrt{y}}.$$

### 2.4.3 Normal distribution example

Assume $X \sim N(5, 3^2)$. Show that $Z = \frac{X-5}{3}$ is a standard normal random variable, i.e., $Z \sim N(0, 1)$.

**Answer:** By using the change of variables and the formula for $f_X(x)$, we have:

$$z = \frac{x-5}{3} \Rightarrow x = 3z + 5 \Rightarrow dx = 3\,dz.$$

Then

$$f_X(x)dx = \frac{1}{3\sqrt{2\pi}}e^{-(x-5)^2/(2\cdot 3^2)}dx = \frac{1}{3\sqrt{2\pi}}e^{-z^2/2}3dz = f_Z(z)dz.$$

Thus $Z$ has the standard normal density.